Université de Metz

Licence de Mathématiques - 3ème année

1er semestre

ANALYSE NUMERIQUE NON-LINEAIRE

par Ralph Chill

Laboratoire de Mathématiques et Applications de Metz

Année 2010/11

Table des matières

CHAPITRE 1

Représentation des nombres

PROPOSITION 1.1. *Soit $\beta \geq 2$ un entier donné (la* base*). Alors tout nombre réel $x \in \mathbb{R}$ s'écrit de la forme*

$$x = (-1)^s \cdot \sum_{j=-\infty}^{k} a_j \beta^j$$

avec $s \in \{0, 1\}$, $k \in \mathbb{Z}$, $a_j \in \mathbb{N}$, $0 \leq a_j \leq \beta - 1$. Si on exige que $a_k \neq 0$, alors la représentation ci-dessus est unique. On écrit aussi

$$x = a_k a_{k-1} \dots a_0, a_{-1} a_{-2} a_{-3} \dots . \tag{1.1}$$

Si $\beta = 10$, alors la représentation (1.1) est exactement la représentation *décimale* de x. Autres bases fréquentes: $\beta = 2$ (représentation binaire), $\beta = 8$ ou $\beta = 16$ (représentation hexadécimale). Pour les ordinateurs, on a typiquement $\beta = 2$.

Dans l'écriture (1.1), déplacer la virgule d'une position à droite (ou à gauche) correspond à une multiplication (ou une division) par β. Tout nombre $x \in \mathbb{R}$ admet donc une unique représentation sous la forme

$$x = (-1)^s a_k, a_{k-1} a_{k-2} a_{k-3} \cdots \cdot \beta^k \tag{1.2}$$

avec $s \in \{0, 1\}$, $k \in \mathbb{Z}$, $a_j \in \mathbb{N}$, $0 \leq a_j \leq \beta - 1$, $a_k \neq 0$. C'est cette représentation *scientifique* (ou: représentation *en virgule flottante*) que l'on va utiliser dans la suite. Dans cette représentation on appelle s ou $(-1)^s$ le *signe*, la suite $a_k, a_{k-1} a_{k-2} a_{k-3} \dots$ la *mantisse* et $k \in \mathbb{Z}$ l'*exposant* de x. Ces trois variables caractérisent le nombre réel x.

0.1. Représentation binaire. Nombres machine. Dans la représentation binaire (c.à.d. $\beta = 2$) on a

$$\begin{aligned} &s \in \{0,\ 1\}, \\ &k \in \mathbb{Z}, \\ &a_j \in \{0,\ 1\} \quad (j \leq k), \text{ et} \\ &a_k = 1. \end{aligned}$$

Comme la mémoire d'un ordinateur est limité, on ne peut coder qu'un nombre fini de nombres réels. Ainsi, la mantisse d'un *nombre machine* est toujours une suite finie, et l'ensemble des exposants k est un ensemble fini. Par exemple, avec *b bits* (*binary*

digits) on ne peut coder que 2^b nombres réels. Pour la compatibilité entre ordinateurs, on utilise le standard IEEE (Institute of Electrical and Electronics Engineers) suivant.

0.1.1. *Simple précision (32 bits).* Avec 32 bits on peut coder $2^{32} \approx 4,29$ milliards de nombres réels. Selon le standard IEEE on utilise

1 bit	pour le signe
23 bits	pour la mantisse, et
8 bits	pour l'exposant k qui varie entre -126 et 127.

On remarque que pour un nombre réel non-nul, on a toujours $a_k = 1$ et qu'il suffit donc de coder les 23 bits de la mantisse qui suivent aprés la virgule. L'exposant est décalé de $2^7 - 1 = 127$. Pour les nombres machine non-nuls on a donc

$$\underbrace{S}_{\text{signe (1 bit)}} \; 1, \underbrace{MMM\ldots MMM}_{\text{mantisse (23 bits)}} \underbrace{EEEEEEEE}_{\text{exposant (8 bits)}}.$$

L'exposant 00000000 avec la mantisse 00000000000000000000000 représente le nombre $x = 0$. L'exposant 11111111 avec la mantisse 00000000000000000000000 représente l'infini. Le plus petit nombre machine positif, non-nul est

$$\begin{aligned} &0 \; 1,00000000000000000000000 \; 00000001 \\ =&1 \cdot 2^{-126} \\ \approx&1,175 \cdot 10^{-38}. \end{aligned}$$

On appelle *précision machine* le plus petit nombre machine positif eps tel que $1 \boxplus \text{eps} > 1$ (addition en virgule flottante, c.à.d. le résultat $1 \boxplus \text{eps}$ est de nouveau un nombre machine). En simple précision, on a

$$\text{eps} = 2^{-23} \approx 1,2 \cdot 10^{-7}.$$

En général, on a

$$x \cdot y = (x \odot y)(1 + r)$$

pour deux nombres machine x, y. Ici, $x \cdot y$ est la multiplication exacte, $x \odot y$ est la multiplication en virgule flottante, et l'*erreur relatif* r vérifie $|r| \leq \text{eps}$.

0.1.2. *Double précision (64 bits).* Avec 64 bits on peut coder $2^{64} \approx 1,8 \cdot 10^{19}$ nombres réels. Selon le standard IEEE on utilise

1 bit	pour le signe
52 bits	pour la mantisse, et
11 bits	pour l'exposant k qui varie entre -1022 et 1023.

On remarque qu'on a toujours $a_k = 1$ et qu'il suffit donc de coder les 52 bits de la mantisse qui suivent aprés la virgule. L'exposant est décalé de $2^{11} - 1 = 1023$.

$$\underbrace{S}_{\text{signe (1 bit)}} \; 1, \underbrace{MMM\ldots MMM}_{\text{mantisse (52 bits)}} \underbrace{EEEEEEEEEEE}_{\text{exposant (11 bits)}}.$$

Le plus petit nombre machine positif, non-nulle est donc

$$\begin{aligned} &0\ 1,00000000000000000000\ldots\ 00000000001\\ =&1\cdot 2^{-1022}. \end{aligned}$$

En double précision, on a

$$\text{eps} = 2^{-52} \approx 2,2\cdot 10^{-16}.$$

CHAPITRE 2

Interpolation et approximation

1. Interpolation selon Lagrange

On considère le problème d'interpolation polynômiale suivant: étant donné

$$x_0, \ldots, x_n \in \mathbb{R} \quad \text{(les noeuds) et}$$
$$y_0, \ldots, y_n \in \mathbb{R} \quad \text{(les abscisses),}$$

on cherche un polynôme $p : \mathbb{R} \to \mathbb{R}$ de degré $\leq n$ tel que

$$p(x_i) = y_i \text{ pour tout } 0 \leq i \leq n.$$

Bien sur, au lieu de chercher un polynôme classique de degré $\leq n$, on peut aussi chercher un polynôme trigonométrique, une fonction spline (une fonction de classe C^2 qui est un polynôme par morceaux), un élément fini ... l'important étant qu'on cherche une fonction interpolante dans un espace vectoriel donné et de dimension finie. Par exemple, l'espace

$$\mathbb{R}_n[X] = \{p : \mathbb{R} \to \mathbb{R} : p \text{ est polynôme de degré } \leq n\}$$

des polynômes de degré $\leq n$ est un espace de dimension $n + 1$. Une base vectorielle est constitué des monômes x^i $(0 \leq i \leq n)$ et tout polynôme $p \in \mathbb{R}_n[X]$ admet une représentation unique de la forme

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

avec $a_i \in \mathbb{R}$.

Proposition 2.1. *Etant donné des noeuds $x_0 < \cdots < x_n$, et étant donné dea abscisses $y_0, \ldots, y_n \in \mathbb{R}$, il existe un et un seul polynôme $p_n \in \mathbb{R}_n[X]$ tel que $p(x_i) = y_i$ pour tout $0 \leq i \leq n$. En d'autres mots, le problème d'interpolation polynômiale admet une unique solution.*

Première démonstration de la Proposition 2.1 (compliquée?) Si on représente les polynômes comme combinaisons linéaires de la base vectorielle des monômes, alors le problème d'interpolation est le problème de montrer existence et unicité d'une famille de coefficients $a_0, \ldots, a_n$ tel que $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ vérifie $p(x_i) = y_i$ $(0 \leq i \leq n)$. Ceci est équivalent à résoudre le système linéaire

$$\begin{aligned} a_0 + a_1 x_0 + \cdots + a_n x_0^n &= y_0 \\ &\vdots \quad \vdots \\ a_0 + a_1 x_n + \cdots + a_n x_n^n &= y_n, \end{aligned}$$

c.à.d. le système

$$\begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Ce système admet une unique solution $(a_0, a_1, \ldots, a_n)$ si et seulement si la *matrice de Vandermonde*

$$A = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}$$

est inversible. Autrement dit, le problème d'interpolation admet une unique solution si et seulement si $\det A \neq 0$.

Pour calculer le determinant $\det A$, on constate que, en développant par exemple par la dernière ligne, que $\det A = f(x_0, \ldots, x_n)$ est un polynôme de degré $\leq n$ en la variable x_n. En remplaçant x_n par x_j (avec $0 \leq j \leq n-1$) on voit que $f(x_0, \ldots, x_{n-1}, x_j) = 0$, c.à.d. chaque noeud x_j $(0 \leq j \leq n-1)$ est une racine de ce polynôme. Ainsi

$$f(x_0, \cdot, x_n) = C \cdot (x_n - x_0) \cdot \cdots \cdot (x_n - x_{n-1})$$

pour une constante $C \in \mathbb{R}$. Cette constante est le coefficient principal du polynôme $f(x_0, \ldots, x_n)$, et en revenant au développement de $\det A$ par la dernière ligne, on voit que $C = f(x_0, \ldots, x_{n-1})$, le déterminant de la matrice de Vandermonde associé aux n noeuds $x_0, \ldots, x_{n-1}$. Par récurrence, on trouve alors

$$\det A = \prod_{0 \leq i < j \leq n} (x_j - x_i),$$

et comme les noeuds sont tous distincts, $\det A \neq 0$. □

Deuxième démonstration de la Proposition 2.1 (plus élégante?) L'idée principale de la démonstration est d'utiliser une autre base vectorielle de l'espace $\mathbb{R}_n[X]$. Pour tout $i \in \{0, \ldots, n\}$ on définit le polynôme

$$\ell_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} \qquad (x \in \mathbb{R}).$$

Alors ℓ_i est un polynôme de degré n avec la propriété importante que

$$\ell_i(x_j) = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases}$$

En conséquence, la famille $(\ell_i)_{0 \leq i \leq n}$ est linéairement indépendante. Mais comme la dimension de $\mathbb{R}_n[X]$ est $n+1$, c.à.d. exactement le nombre d'élement de cette famille, on obtient que $(\ell_i)_{0 \leq i \leq n}$ est une base vectorielle de $\mathbb{R}_n[X]$.

Existence d'un polynôme d'interpolation. Il suffit maintenant de poser

$$p_n(x) = \sum_{i=0}^{n} y_i \, \ell_i(x) \qquad (x \in \mathbb{R}). \tag{2.1}$$

Comme combinaison linéaire des ℓ_i, p_n est un polynôme de degré $\leq n$. En plus,

$$p_n(x_j) = \sum_{i=0}^{n} y_i \, \ell_i(x_j) = y_j \qquad \text{pour tout } 0 \leq j \leq n.$$

Donc, le polynôme p_n *est* un polynôme d'interpolation. D'où l'existence d'un polynôme d'interpolation.

Unicité. Soit $p \in \mathbb{R}_n[X]$ un polynôme d'interpolation. Alors, comme $(\ell_i)_{0 \leq i \leq n}$ est une base vectorielle de $\mathbb{R}_n[X]$,

$$p(x) = \sum_{i=0}^{n} z_i \ell_i(x) \qquad (x \in \mathbb{R})$$

pour des $z_i \in \mathbb{R}$. Alors, quelque soit $j \in \{0, \ldots, n\}$,

$$\begin{aligned} z_j &= \sum_{i=0}^{n} z_i \, \ell_i(x_j) && \text{(propriété des } \ell_i\text{)} \\ &= p(x_j) \\ &= y_j && \text{(p est polynôme d'interpolation),} \end{aligned}$$

et donc

$$p(x) = \sum_{i=0}^{n} z_i \, \ell_i(x) = \sum_{i=0}^{n} y_i \, \ell_i(x) = p_n(x),$$

où p_n est le polynôme de (2.1). D'où l'unicité. □

On appelle le polynôme ℓ_i le *i-iéme polynôme de Lagrange*. D'après la Proposition 2.1 nous pouvons (et nous allons) parler *du* polynôme d'interpolation. La représentation (2.1) est la représentation de Lagrange du polynôme d'interpolation associé aux noeuds x_i et aux abscisses y_i. Cette représentation est simple, mais il faut noter que pour calculer le i-iéme polynôme de Lagrange il faut faire $2n$ additions, n divisions et n multiplications. Eventuellement on divise par des nombres petits si les noeuds sont proches l'un de l'autre; ceci peut engendrer des erreurs d'arrondis importants. Dans la suite nous verrons d'autres représentations de ce polynôme, ou au moins d'autres algorithmes pour calculer $p_n(x)$ en un point x donné.

2. Etude de l'erreur d'interpolation

On suppose dans cette section que les abscisses $y_i = f(x_i)$ pour une fonction $f : [a, b] \to \mathbb{R}$ donnée, et que les noeuds vérifient $a \leq x_0 < \cdots < x_n \leq b$.

PROPOSITION 2.2. *Soit p_n le polynôme d'interpolation de Lagrange associé aux noeuds $a \le x_0 < \cdots < x_n \le b$ et aux abscisses $f(x_0)$, ..., $f(x_n)$, où $f \in C^{n+1}([a,b])$ est une fonction donnée. Alors pour tout $x \in [a,b]$ il existe $\xi_x \in]a,b[$ tel que*

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{i=0}^{n}(x - x_i).$$

DÉMONSTRATION. Par définition du polynôme d'interpolation, $f(x) - p_n(x) = 0$ si $x = x_i$ pour un $i \in \{0, \ldots, n\}$. On peut alors supposer que $x \neq x_i$ pour tout $i \in \{0, \ldots, n\}$. On pose

$$Q(t) := p_n(t) - f(t) + \frac{f(x) - p_n(x)}{\prod_{i=0}^{n}(x - x_i)} \prod_{i=0}^{n}(t - x_i) \qquad (t \in [a,b]).$$

Alors $Q \in C^{n+1}([a,b])$ et

$$Q(x_i) = 0 \quad \text{pour tout } i, \text{ et } Q(x) = 0.$$

Le théorème de Rolle implique que la dérivée Q' s'annulle en $n+1$ points distincts dans l'intervalle $]a,b[$. En appliquant le théorème de Rolle encore une fois, on voit que Q'' s'annulle en n points distincts dans l'intervalle $]a,b[$, et finalement, qu'il existe $\xi_x \in]a,b[$ tel que $Q^{(n+1)}(\xi_x) = 0$. La proposition s'en déduit facilement. □

COROLLAIRE 2.3. *Soit $\|f^{(n+1)}\|_\infty := \sup_{x\in[a,b]} |f^{(n+1)}(x)|$. Alors*

$$|f(x) - p_n(x)| \le \frac{\|f^{(n+1)}\|_\infty}{(n+1)!} \prod_{i=0}^{n}(x - x_i) \qquad (x \in [a,b]).$$

La proposition et le corollaire montrent que l'erreur d'interpolation dépend aussi du choix des noeuds x_i. En fait,

$$|f(x) - p_n(x)| \le \frac{\|f^{(n+1)}\|_\infty}{(n+1)!} \sup_{t\in[a,b]} \Big| \prod_{i=0}^{n}(t - x_i) \Big| \qquad (x \in [a,b]).$$

On verra plus loin qu'un choix optimal est le choix des noeuds de Tchebychev, c.à.d., sur l'intervalle $[a,b] = [-1,1]$, le choix des noeuds

$$x_i = \cos \frac{(2i+1)\pi}{2n+2} \qquad (0 \le i \le n).$$

3. L'algorithme de Newton

L'algorithme de Newton est une méthode efficace pour calculer le polynôme d'interpolation associé à des noeuds x_0, ..., x_n donnés et à des abscisses y_0, ..., y_n donnés (on suppose toujours que $x_i \neq x_j$ si $i \neq j$). Il est particulierement pratique dans les situations où on veut rajouter des noeuds. L'idée est de représenter le polynôme d'interpolation dans la forme

$$p(x) = c_0 + c_1 (x - x_0) + \cdots + c_n (x - x_0) \cdot \cdots \cdot (x - x_{n-1}) \tag{2.2}$$

avec des coefficients c_0, ..., c_n à determiner. On remarque ici que les polynômes 1, $x - x_0$, ..., $(x - x_0) \cdot \cdots \cdot (x - x_{n-1})$ forment une base de l'espace $\mathbb{R}_n[X]$ (exercice!)

et que les coefficients $c_0, \ldots, c_n$ existent et sont uniques.

Dans la suite, pour tout $i,\ j \in \{0, \ldots, n\},\ j \le i$, on note

$$\begin{aligned} p_{i,j} = {} & \text{le polynôme d'interpolation de degré } \le j \\ & \text{associé aux noeuds } x_{i-j},\ \ldots,\ x_i \\ & \text{et aux abscisses } y_{i-j},\ \ldots, y_i. \end{aligned}$$

Le polynôme que l'on recherche est le polynôme $p_{n,n}$. On note ensuite

$$\begin{aligned} c_{i,j} = {} & \text{le coefficient principal du polynôme } p_{i,j}, \text{c.à.d. le coefficient} \\ & \text{correspondant au monôme } x^j \text{ dans la représentation} \\ & p_{i,j}(x) = c_{i,j}x^j + \text{ polynôme de degré } < j. \end{aligned}$$

On rappelle que, étant donné des noeuds $x_0, \ldots, x_n$ et des abscisses $y_0, \ldots, y_n$, le polynôme d'interpolation p de degré $\le n$ existe et est unique (Proposition 2.1). Ainsi, le coefficient principal est déterminé de manière unique en fonction des noeuds et des abscisses.

Remarque 2.4. Le coefficient principal du polynôme d'interpolation correspondant aux noeuds $x_0, \ldots, x_n$ et aux abscisses $y_0, \ldots, y_n$ est dans la littérature aussi noté $f[x_0, \ldots, x_n]$ (la notation avec le f s'explique si les abscisses $y_0, \ldots, y_n$ sont de la forme $f(x_0), \ldots, f(x_n)$ pour une fonction f donnée). Avec cette notation, on aurait

$$c_{i,j} = f[x_{i-j}, \ldots, x_i].$$

Lemme 2.5. *Soit $p = p_{n,n}$ le polynôme d'interpolation correspondant aux noeuds $x_0, \ldots, x_n$ et aux noeuds $y_0, \ldots, y_n$. On représente p dans la forme* (2.2)*, c.à.d.*

$$p(x) = c_0 + c_1\,(x - x_0) + \cdots + c_n\,(x - x_0) \cdot \cdots \cdot (x - x_{n-1}) \qquad (x \in \mathbb{R}).$$

Alors $c_i = c_{i,i}$ (le coefficient principal du polynôme $p_{i,i}$).

Démonstration. Exercice. □

Afin de calculer des coefficients $c_{i,j}$, on note d'abord que

$$c_{i,0} = y_i \qquad (0 \le i \le n). \tag{2.3}$$

En fait, $c_{i,0}$ est, par définition, le coefficient principal du polynôme $p_{i,0}$ de degré ≤ 0 (c.à.d. $p_{i,0}$ est constant!) tel que $p_{i,0}(x_i) = y_i$. Ainsi, $p_{i,0}(x) = y_i = y_i \cdot x^0$, d'où (2.3). Afin de calculer les autres coefficients $c_{i,j}$ avec $j \ge 1$ on utilise le lemme et le corollaire suivant.

Lemme 2.6 (Différences divisées). *Pour tout $i,\ j \in \{1, \ldots, n\}$ avec $i \ge j$ on a*

$$p_{i,j}(x) = \frac{p_{i,j-1}(x)(x - x_{i-j}) - p_{i-1,j-1}(x)(x - x_i)}{x_i - x_{i-j}} \qquad (x \in \mathbb{R}).$$

Démonstration. Soit $q(x) := \frac{p_{i,j-1}(x)(x-x_{i-j})-p_{i-1,j-1}(x)(x-x_i)}{x_i-x_{i-j}}$ $(x \in \mathbb{R})$. Par définition, $p_{i,j-1}$ et $p_{i-1,j-1}$ sont des polynômes de degré $\leq j-1$. Ainsi, q est un polynôme de degré $\leq j$. Ensuite, on calcul

$$q(x_{i-j}) = \frac{0 - p_{i-1,j-1}(x_{i-j})(x_{i-j} - x_i)}{x_i - x_{i-j}} = y_{i-j},$$

car $p_{i-1,j-1}(x_{i-j}) = y_{i-j}$ par définition de $p_{i-1,j-1}$. De manière similaire

$$q(x_i) = \frac{p_{i,j-1}(x_i)(x_i - x_{i-j}) - 0}{x_i - x_{i-j}} = y_i,$$

car $p_{i,j-1}(x_i) = y_i$ par définition de $p_{i,j-1}$. Finalement, pour tout k avec $i-j < k < i$ on a $p_{i,j-1}(x_k) = y_k$ et $p_{i-1,j-1}(x_k) = y_k$. Ainsi,

$$\begin{aligned} q(x_k) &= \frac{p_{i,j-1}(x_k)(x_k - x_{i-j}) - p_{i-1,j-1}(x_k)(x_k - x_i)}{x_i - x_{i-j}} \\ &= \frac{y_k(x_k - x_{i-j}) - y_k(x_k - x_i)}{x_i - x_{i-j}} \\ &= y_k. \end{aligned}$$

On a donc montré que q est un polynôme d'interpolation de degré $\leq j$ associé aux noeuds $x_{i-j}, \ldots, x_i$ et aux abscisses $y_{i-j}, \ldots, y_i$. Par unicité du polynôme d'interpolation, $q = p_{i,j}$. □

Corollaire 2.7 (Différences divisées). *Pour tout $i,\ j \in \{1, \ldots, n\}$ avec $i \geq j$ on a*

$$c_{i,j} = \frac{c_{i,j-1} - c_{i-1,j-1}}{x_i - x_{i-j}}. \tag{2.4}$$

Démonstration. Conséquence directe du Lemme 2.6 et de la définition des $c_{i,j}$. □

D'après les équations (2.3) et (2.4), on obtient le schéma suivant pour calculer successivement les $c_{i,j}$ (algorithme de Newton):

$$\begin{array}{c|ccccccccc} x_0 & c_{0,0} = y_0 & & & & & & & & \\ & & \searrow & & & & & & & \\ x_1 & c_{1,0} = y_1 & \to & c_{1,1} & & & & & & \\ & & \searrow & & \searrow & & & & & \\ x_2 & c_{2,0} = y_2 & \to & c_{2,1} & \to & c_{2,2} & & & & \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & & & \\ x_{n-1} & c_{n-1,0} = y_{n-1} & \to & c_{n-1,1} & \to & c_{n-1,2} & \ldots & c_{n-1,n-1} & & \\ & & \searrow & & \searrow & & & & \searrow & \\ x_n & c_{n,0} = y_n & \to & c_{n,1} & \to & c_{n,2} & \ldots & c_{n,n-1} & \to & c_{n,n} \end{array}$$

On a finalement

$$p_{n,n}(x) = c_{0,0} + c_{1,1}\,(x - x_0) + \cdots + c_{n,n}\,(x - x_0) \cdot \cdots \cdot (x - x_{n-1}) \qquad (x \in \mathbb{R}).$$

Exemple 2.8. On cherche le polynôme d'interpolation de degré ≤ 3 associé aux noeuds

$$x_0 = -1, \quad x_1 = 0, \quad x_2 = 1, \quad x_3 = 3$$

et aux abscisses

$$y_0 = 1, \quad y_1 = -1, \quad y_2 = 0, \quad y_3 = 2.$$

Avec l'algorithme de Newton on trouve

$$\begin{array}{l|llllllll}
x_0 = -1 & c_{0,0} = y_0 = 1 & & & & & & & \\
 & & \searrow & & & & & & \\
x_1 = 0 & c_{1,0} = y_1 = -1 & \to & c_{1,1} = -2 & & & & & \\
 & & \searrow & & \searrow & & & & \\
x_2 = 1 & c_{2,0} = y_2 = 0 & \to & c_{2,1} = 1 & \to & c_{2,2} = \frac{3}{2} & & \\
 & & \searrow & & \searrow & & \searrow & & \\
x_3 = 3 & c_{3,0} = y_3 = 2 & \to & c_{3,1} = 1 & \to & c_{3,2} = 0 & \to & c_{3,3} = -\frac{3}{8}
\end{array}$$

Ici, on a calculé successivement, colonne par colonne,

$$\begin{aligned}
c_{1,1} &= \frac{c_{1,0} - c_{0,0}}{x_1 - x_0} = \frac{-1-1}{0-(-1)} = -2, \\
c_{2,1} &= \frac{c_{2,0} - c_{1,0}}{x_2 - x_1} = \frac{0-(-1)}{1-0} = 1, \\
c_{3,1} &= \frac{c_{3,0} - c_{2,0}}{x_3 - x_2} = \frac{2-0}{3-1} = 1, \\
c_{2,2} &= \frac{c_{2,1} - c_{1,1}}{x_2 - x_0} = \frac{1-(-2)}{1-(-1)} = \frac{3}{2}, \\
c_{3,2} &= \frac{c_{3,1} - c_{2,1}}{x_3 - x_1} = \frac{1-1}{3-0} = 0 \\
c_{3,3} &= \frac{c_{3,2} - c_{2,2}}{x_3 - x_0} = \frac{0-\frac{3}{2}}{3-(-1)} = -\frac{3}{8}.
\end{aligned}$$

On trouve alors

$$\begin{aligned}
p(x) &= 1 - 2\,(x-(-1)) + \frac{3}{2}\,(x-(-1))\,(x-0) - \frac{3}{8}\,(x-(-1))\,(x-0)\,(x-1) \\
&= 1 - 2\,(x+1) + \frac{3}{2}\,(x+1)\,x - \frac{3}{8}\,(x+1)\,x\,(x-1) \\
&= -1 - \frac{1}{8}\,x + \frac{3}{2}\,x^2 - \frac{3}{8}\,x^3,
\end{aligned}$$

et il est facile a vérifier que c'est effectivement le polynôme d'interpolation recherché.

4. Approximation hilbertienne

Soit $I \subseteq \mathbb{R}$ un intervalle et soit $w : I \to \mathbb{R}$ une fonction continue positive (une *fonction poids*). On considère l'espace

$$C_w(I) := \{f \in C(I) : \int_I |f(x)|^2 w(x)\,dx < \infty\},$$

où $C(I)$ est l'espace de toutes les fonctions continues de I dans $\mathbb{R}$. L'espace $C_w(I)$ sera muni du produit scalaire $\langle \cdot, \cdot \rangle_w$ donné par

$$\langle f, g \rangle_w := \int_I f(x)g(x)w(x)\, dx \qquad (f,\ g \in C_w(I)).$$

La norme associée à ce produit scalaire est la norme $\| \cdot \|_w$ donnée par

$$\|f\|_w := \sqrt{\langle f, f \rangle_w} = \left(\int_I |f(x)|^2 w(x)\, dx \right)^{\frac{1}{2}}.$$

Soit $F \subseteq C_w(I)$ un sous-espace vectoriel de dimension finie, et soit $f \in C_w(I)$ une fonction donnée. Le probléme de la meilleure approximation de f par un élément de F est le problème de trouver une fonction $p \in F$ telle que

$$\|f - p\|_w = \inf_{q \in F} \|f - q\|_w,$$

c.à.d. de trouver une fonction $p \in F$ telle que la distance $\|f - p\|_w$ soit minimale parmi toutes les distances $\|f - q\|_w$ $(q \in F)$.

4.1. Solution théorique. La proposition suivante montre que le problème de la meilleure approximation de f admet une unique solution.

Proposition 2.9. *Soit E un espace vectoriel réel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Soit $f \in E$ et soit $F \subseteq E$ un sous-espace vectoriel de dimension finie. Alors il existe un unique élément $p \in F$ telle que $\|f - p\| = \inf_{q \in F} \|f - q\|$ (ici: $\|g\| := \sqrt{\langle g, g \rangle}$). Cet élément p est appelé la* projection orthogonale *de f sur F. Il est caractérisée par le fait que*

$$\langle f - p, q \rangle_w = 0 \text{ pour tout } q \in F. \tag{2.5}$$

Démonstration. On démontre d'abord que $p \in F$ est une meilleure approximation de f dans F si et seulement si la condition (2.5) est vérifiée. En fait, $p \in F$ est une meilleure approximation de f dans F si et seulement si $p \in F$ et

$$\|f - p\|^2 \leq \|f - q\|^2 \text{ pour tout } q \in F,$$

ce qui est équivalent à

$$-2\langle f, p \rangle + \|p\|^2 \leq -2\langle f, q \rangle + \|q\|^2 \text{ pour tout } q \in F.$$

En remplaçant $q \in F$ par $p + q \in F$, ceci est équivalent à

$$\langle f - p, q \rangle \leq \|q\|^2 \text{ pour tout } q \in F.$$

Ici, on remplace q par tq $(t \in \mathbb{R},\ t > 0)$, on divise par t, et on fait t tendre vers 0 pour obtenir

$$\langle f - p, q \rangle \leq 0 \text{ pour tout } q \in F.$$

Finalement, en remplaçant q par $-q$, on obtient que $p \in F$ est une meilleure approximation de f dans F si et seulement si la condition (2.5) est vérifiée.

Existence et unicité d'une meilleure approximation. Soit $(p_i)_{0 \leq i \leq n}$ une base vectorielle de l'espace F (on suppose donc qu'il est de dimension $n + 1$. Alors tout élément $p \in F$ s'écrit de la forme $p = \sum_{i=0}^{n} \lambda_i p_i$ avec des coefficients $\lambda_i \in \mathbb{R}$. En

plus, come (p_i) est une base vectorielle de F, la condition (2.5) est équivalente à la condition

$$\langle f, p_i\rangle = \sum_{j=0}^{n} \lambda_j \langle p_j, p_i\rangle \text{ pour tout } 0 \le i \le n, \tag{2.6}$$

ce qui est un système linéaire pour le vecteur $(\lambda_0, \dots, \lambda_n)$. Ce système linéaire admet une unique solution $(\lambda_0, \dots, \lambda_n)$ si et seulement si le système homogène

$$\sum_{j=0}^{n} \mu_j \langle p_j, p_i\rangle = 0 \text{ pour tout } 0 \le i \le n$$

admet $(\mu_0, \dots, \mu_n) = (0, \dots, 0)$ comme seule solution. Mais ce système implique que

$$0 = \sum_{i=0}^{n} \mu_i \sum_{j=0}^{n} \mu_j \langle p_j, p_i\rangle$$

$$\| \sum_{i=0}^{n} \mu_i p_i \|^2,$$

et donc $\sum_{i=0} \mu_i p_i = 0$. Comme la famille (p_i) est linéairement indépendante, on trouve $(\mu_0, \dots, \mu_n) = (0, \dots, 0)$. □

La démonstration de la proposition précédente montre que $p = \sum_{i=0}^{n} \lambda_i p_i$ est meilleure approximation de f dans F si et seulement si $(\lambda_0, \dots, \lambda_n)$ est solution de (2.6). Il suffit donc de choisir une base vectorielle (p_i) de F est de résoudre le système linéaire (2.6). On remarque que ce système linéaire devient particulièrement simple à résoudre si (p_i) est une base orthogonale de F, c.à.d.

$$\langle p_j, p_i\rangle = \begin{cases} \|p_i\|^2 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Corollaire 2.10. *Soit $(p_i)_{0\le i\le n}$ une base orthogonale de F. Alors la meilleure approximation p de f dans F est donné par*

$$p = \sum_{i=0}^{n} \frac{\langle f, p_i\rangle}{\|p_i\|^2}\, p_i.$$

4.2. Le procédé d'orthogonalisation de Gram-Schmidt. Soit E un espace vectoriel réel muni d'un produit scalaire $\langle\cdot,\cdot\rangle$. Soit $(e_i)_{i\ge0} \subseteq E$ une famille de vecteurs linéairement indépendante. Alors le procédé suivant permet de construire une famille orthonormale $(p_i)_{i\ge0}$.

Procédé de Gram-Schmidt. On pose d'abord

$$\tilde{e}_0 = e_0 \quad \text{et}$$

$$p_0 = \frac{\tilde{e}_0}{\|\tilde{e}_0\|}.$$

Ensuite, pour tout $i \geq 1$ on définit de manière recursive

$$\tilde{e}_i = e_i - \sum_{j=0}^{i-1} \langle e_i, p_j \rangle p_j \quad \text{et}$$
$$p_i = \frac{\tilde{e}_i}{\|\tilde{e}_i\|}.$$

C'est un exercice de montrer que pour tout $i,\ j \geq 0$ on a

$$\langle p_i, p_j \rangle = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases}$$

c.à.d. que (p_i) est une famille orthonormale. En plus, si (e_i) était une base vectorielle, alors (p_i) est une base vectorielle normale.

4.3. Les polynômes orthogonaux. Soit $I \subseteq \mathbb{R}$ un intervalle quelconque, et soit $w : I \to \mathbb{R}$ une fonction poids (continue, strictement positive). On suppose que, quelque soit $n \in \mathbb{N}$,

$$\int_I x^n\, w(x)\, dx \text{ est absolument convergente.}$$

Dans ce cas, l'espace $C_w(I)$ contient tous les polynômes. En appliquant le procédé de Gram-Schmidt aux monômes $e_n(x) = x^n$ $(n \geq 0)$, on obtient une famille $(p_n)_{n\geq 0}$ de polynômes orthogonales.

Intervalle	Poids $w(x)$	Polynômes
$[-1, 1]$	1	Legendre
$[0, \infty[$	e^{-x}	Laguerre
$\mathbb{R}$	e^{-x^2}	Hermite
$]-1, 1[$	$\frac{1}{\sqrt{1-x^2}}$	Tchebychev .

Par exemple, pour les polynômes de Legendre, on obtient

$$p_0(x) =$$
$$p_1(x) =$$
$$p_2(x) =$$
$$\vdots$$

Lemme 2.11. *Les polynômes orthogonaux vérifient*

(a) *p_n est un polynôme de degré n,*
(b) *la famille $(p_i)_{0\leq i\leq n}$ est une base orthogonale de l'espace $\mathbb{R}_n[X]$,*
(c) *$\int_I p_n(x)q(x)w(x)\, dx = 0$ pour tout polynôme q de degré $< n$.*

Théorème 2.12. *Le polynôme p_n possède n racines réelles. Elles sont simples et contenues dans l'intérieur de l'intervalle I.*

DÉMONSTRATION. On peut écrire

$$p_n(x) = \prod_{i=1}^{k} (x - r_i)^{m_i}\, r(x)$$

où les r_i sont des racines réelles distincts de multiplicité m_i, et le polynôme r n'a pas de racines réelles (en particulier, r a un signe). Notons $r_{i_1}, \ldots, r_{i_s}$ les racines de p_n situées à l'intérieur de l'intervalle I et de multplicités m_{i_α} impairs. Alors le polynôme

$$q(x) = (x - r_{i_1}) \cdot \cdots \cdot (x - r_{i_s})$$

est un polynôme de degré s, toutes les racines de q sont simples et contenues dans l'intérieur de l'intervalle I. En plus, le produit $p_n(x)q(x)$ ne change pas de signe et est non-nulle. Donc $\langle p_n, q\rangle \neq 0$. D'après le lemme précédent (plus précisément, Lemme 2.11 (c), on obtient donc que degré p_n = degré q, c.à.d. que q est un polynôme de degré = n. □

5. Approximation uniforme

Dans cette section, on suppose que $I = [a, b]$ est un intervalle compact. L'espace $C([a, b])$ de toutes les fonctions continues $[a, b] \to \mathbb{R}$ muni de la norme $\|f\|_\infty := \sup_{x\in[a,b]} |f(x)|$ est un espace vectoriel normé. Etant donnée une fonction $f \in C([a, b])$, on considère de nouveau le problème de (meilleure) approximation, mais maintenant avec la norme $\|\cdot\|_\infty$ à la place de la norme hilbertienne $\|\cdot\|_w$ du paragraphe précédent. On rappelle le théorème suivant.

THÉORÈME 2.13 (Weierstrass). *Pour toute fonction $f \in C([a, b])$ il existe une suite (p_n) de polynômes telle que* $\lim_{n\to\infty} \|f - p_n\|_\infty = 0$.

CHAPITRE 3

Intégration et différentiation numérique

1. Intégration approchée

Soit $f : [a, b] \to \mathbb{R}$ une fonction continue donnée, $[a, b] \subseteq \mathbb{R}$ étant un intervalle compacte. On considère le problème de calculer l'intégrale

$$I(f) = \int_a^b f(x)\, dx.$$

En général on cherche à trouver une valeur approchée de l'intégrale sous la forme

$$J(f) = \sum_{i=0}^{n} \lambda_i f(x_i),$$

où $a \leq x_0 < \cdots < x_n \leq b$ sont des noeuds donnés (ou à choisir) et où les coefficients $\lambda_0, \ldots, \lambda_n$ sont donnés (ou à choisir), tous les deux indépendants de la fonction f. Supposons d'abord que les noeuds $a \leq x_0 < \cdots < x_n \leq b$ sont donnés. Comment choisir les coefficients $\lambda_0, \ldots, \lambda_n$ de telle manière que l'erreur

$$E(f) = I(f) - J(f)$$

soit 0 pour une certaine classe de fonctions? Par exemple, comment choisir les coefficients $\lambda_0, \ldots, \lambda_n$ de telle manière que l'erreur $E(f) = 0$ pour tout polynôme de degré inférieur ou égal à n? C.à.d. telle que la formule d'intégration approchée $J(f)$ soit exacte pour tout polynôme de degré inférieur ou égal à n?

Théorème 3.1. *Etant donné des noeuds $a \leq x_0 < \cdots < x_n \leq b$, il existe une et une seule formule d'intégration approchée $J(f) = \sum_{i=0}^{n} \lambda_i f(x_i)$ telle que l'erreur $E(f) = I(f) - J(f) = 0$ pour tout polynôme $f \in \mathbb{R}_n[X]$.*

Démonstration. **Existence.** On pose $\lambda_i = \int_a^b \ell_i(x)\, dx$, oú ℓ_i est le i-éme polynôme d'interpolation de Lagrange associé aux noeuds $x_0, \ldots, x_n$. Alors, pour tout polynôme $f \in \mathbb{R}_n[X]$ on a

$$f(x) = \sum_{i=0}^{n} f(x_i)\, \ell_i(x),$$

car l'interpolation est exacte pour tout polynôme de degré $\leq n$. Donc,

$$\begin{aligned} I(f) &= \int_a^b f(x)\,dx \\ &= \int_a^b \sum_{i=0}^n f(x_i)\,\ell_i(x)\,dx \\ &= \sum_{i=0}^n f(x_i) \int_a^b \ell_i(x)\,dx \\ &= \sum_{i=0}^n \lambda_i f(x_i) \\ &= J(f). \end{aligned}$$

Unicité. Soient $J(f) = \sum_{i=0}^n \lambda_i f(x_i)$ et $J'(f) = \sum_{i=0}^n \lambda_i' f(x_i)$ deux formules d'intégration approchées qui sont exactes pour tous les polynômes $f \in \mathbb{R}_n[X]$, c.à.d.

$$J(f) = \sum_{i=0}^n \lambda_i f(x_i) = I(f) = \sum_{i=0}^n \lambda_i' f(x_i) \text{ pour tout } f \in \mathbb{R}_n[X].$$

Alors

$$\sum_{i=0}^n (\lambda_i - \lambda_i') f(x_i) = 0 \text{ pour tout } f \in \mathbb{R}_n[X].$$

En particulier, en choisissant $f = \ell_i$ le i-ième polynôme d'interpolation de Lagrange (qui a la propriété que $\ell_i(x_j) = \delta_{ij}$ pour tout $1 \leq i, j \leq n$), on trouve que

$$\lambda_i = \lambda_i',$$

d'où l'unicité. □

Définition 3.2. Une formule d'intégration approchée est dite d'ordre m si $I(f) = J(f)$ pour tout polynôme $f \in \mathbb{R}_m[X]$ et s'il existe un polynôme $f \in \mathbb{R}_{m+1}[X]$ tel que $I(f) \neq J(f)$.

D'après le Théorème 3.1, étant donné des noeuds $a \leq x_0 < \cdots < x_n \leq b$, il existe une et une seule formule d'intégration approchée d'ordre $m \geq n$. On verra que l'ordre m peut être strictement supérieur à n. Par contre, on remarque d'abord que l'ordre est toujours inférieur ou égal à $2n + 1$.

Lemme 3.3. *Soient $a \leq x_0 < \cdots < x_n \leq b$ des noeuds donnés et soit $J(f) = \sum_{i=0}^n \lambda_i f(x_i)$ une méthode d'intégration approchée. Alors l'ordre de cette méthode est inférieur ou égal à $2n + 1$.*

Démonstration. Soit

$$f(x) = \prod_{i=0}^n (x - x_i)^2.$$

Alors f est un polynôme positif, de degré $2n+2$, tel que $f(x_i) = 0$ pour tout $0 \leq i \leq n$. Donc

$$J(f) = \sum_{i=0}^{n} \lambda_i f(x_i) = 0 < I(f).$$

Ainsi, l'ordre de J est inférieur ou égal à $2n + 1$. □

Exemple 3.4 (Méthode des rectangles). On prend $n = 0$ et $x_0 = \frac{a+b}{2}$ (un seul noeud, au milieu de l'intervalle $[a, b]$). Alors

$$\lambda_0 = \int_a^b \ell_0(x)\, dx = \int_a^b dx = b - a,$$

et donc

$$J(f) = (b - a) f(\frac{a + b}{2}).$$

Cette formule d'intégration approchée est exacte pour les polynômes constants $f \in \mathbb{R}_0[X]$ (Théorème 3.1), mais on a aussi

$$J(x) = (b - a)\frac{a + b}{2} = \frac{b^2 - a^2}{2} = I(x).$$

Par contre,

$$J(x^2) = (b - a)\frac{(a + b)^2}{4} \neq \frac{b^3 - a^3}{3} = I(x^2).$$

Donc, l'ordre de cette formule d'intégration approchée est $m = 1$.

Exemple 3.5 (Méthode des trapézes). On prend $n = 1$, $x_0 = a$, $x_1 = b$ (deux noeuds, aux extrémités de l'intervalle $[a, b]$). On a

$$\lambda_0 = \int_a^b \ell_0(x)\, dx = \int_a^b \frac{x - b}{a - b}\, dx = \frac{b - a}{2}$$

et

$$\lambda_1 = \int_a^b \ell_1(x)\, dx = \int_a^b \frac{x - a}{b - a}\, dx = \frac{b - a}{2}.$$

Donc,

$$J(f) = \frac{b - a}{2}\,(f(a) + f(b)).$$

Cette formule d'intégration approchée est exacte pour les polynômes $f \in \mathbb{R}_1[X]$ (Théorème 3.1), mais

$$J(x^2) \neq I(x^2).$$

L'ordre de cette formule d'intégration approchée est $m = 1$.

Exemple 3.6 (Méthode de Simpson). On prend $n = 2$, $x_0 = a$, $x_1 = \frac{a+b}{2}$, et $x_2 = b$ (trois noeuds, aux extrémités et au milieu de l'intervalle $[a, b]$). On a

$$\lambda_0 = \int_a^b \ell_0(x)\, dx = \frac{b-a}{6},$$
$$\lambda_1 = \int_a^b \ell_1(x)\, dx = 4\,\frac{b-a}{6},$$
$$\lambda_2 = \int_a^b \ell_2(x)\, dx = \frac{b-a}{6}.$$

Donc

$$J(f) = \frac{b-a}{6}\,(f(a) + 4\,f(\frac{a+b}{2}) + f(b)).$$

L'ordre de cette formule d'intégration approchée est $m = 3$.

2. Etude de l'erreur d'intégration approchée

On rappelle la version suivante du théorème de Taylor, et aussi le théorème de la moyenne.

Théorème 3.7 (Taylor). *Soit $f \in C^{n+1}([a, b])$. Alors pour tout $x \in [a, b]$ on a*

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(x-t)_+^n}{n!} f^{(n+1)}(t)\, dt,$$

où

$$t_+ = \begin{cases} t & si\ t \geq 0, \\ 0 & si\ t < 0. \end{cases}$$

Théorème 3.8 (Théorème de la moyenne). *Soient f, $g : [a, b] \to \mathbb{R}$ deux fonctions continues. On suppose que $g(x) \geq 0$ pour tout $x \in [a, b]$. Alors il existe un $\xi \in [a, b]$ tel que*

$$\int_a^b f(x)g(x)\, dx = f(\xi) \int_a^b g(x)\, dx.$$

Théorème 3.9. *Soient $a \leq x_0 < \cdots < x_n \leq b$ des noeuds donnés et soit $J(f) = \sum_{i=0}^{n} \lambda_i f(x_i)$ une méthode d'intégration approchée d'ordre $m \geq n$. Soit*

$$K(t) = \int_a^b (x-t)_+^m\, dx - \sum_{i=0}^{n} \lambda_i (x_i - t)_+^m \qquad (t \in [a, b]).$$

Alors, pour toute fonction $f \in C^{(m+1)}([a, b])$ on a

$$E(f) = I(f) - J(f) = \int_a^b \frac{K(t)}{m!} f^{(m+1)}(t)\, dt,$$

et si le noyau K est positif, alors

$$E(f) = f^{(m+1)}(\xi) \int_a^b \frac{K(t)}{m!}\, dt \text{ pour un } \xi \in [a,b].$$

Démonstration. D'après le Théorème de Taylor (Théorème 3.7), on a

$$f(x) = \sum_{k=0}^{m} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(x-t)_+^m}{m!} f^{(m+1)}(t)\, dt.$$

Comme J est d'ordre m, on obtient donc

$$\begin{aligned}
E(f) &= E(x \mapsto \int_a^b \frac{(x-t)_+^m}{m!} f^{(m+1)}(t)\, dt) \\
&= \int_a^b \int_a^b \frac{(x-t)_+^m}{m!} f^{(m+1)}(t)\, dt\, dx - \sum_{i=0}^{n} \lambda_i \int_a^b \frac{(x_i-t)_+^m}{m!} f^{(m+1)}(t)\, dt \\
&= \int_a^b \left(\int_a^b (x-t)_+^m\, dx - \sum_{i=0}^{n} \lambda_i (x_i-t)_+^m) \right) f^{(m+1)}(t)\, dt \\
&= \int_a^b \frac{K(t)}{m!} f^{(m+1)}(t)\, dt.
\end{aligned}$$

La deuxième représentation de l'erreur est une conséquence directe de cette première représentation et du théorème de la moyenne (Théorème 3.8). □

Exemple 3.10 (L'erreur dans la méthode des rectangles). Cette méthode est d'ordre 1. Donc,

$$\begin{aligned}
E(f) &= \int_a^b \left(\int_a^b (x-t)_+\, dx - (b-a)(\frac{a+b}{2} - t)_+) \right) f''(t)\, dt \\
&= \int_a^b K(t) f''(t)\, dt
\end{aligned}$$

avec

$$K(t) = \begin{cases} \frac{(t-a)^2}{2} & \text{si } t \le \frac{a+b}{2}, \\ \frac{(t-b)^2}{2} & \text{si } t \ge \frac{a+b}{2}. \end{cases}$$

Ainsi, comme le noyau K est positif et par le théorème de la moyenne,

$$E(f) = f''(\xi) \frac{(b-a)^3}{24} \text{ pour un certain } \xi \in [a,b].$$

Exemple 3.11 (L'erreur dans la méthode de Simpson). Cette méthode est d'ordre 3. Donc,

$$E(f) = \int_a^b \frac{K(t)}{3!} f^{(4)}(t)\, dt$$

avec

$$K(t) = \int_a^b (x-t)_+^3 \, dx - \frac{b-a}{6}((a-t)_+^3 + 4\,(\frac{a+b}{2} - t)_+^3 + (b-t)_+^3)$$
$$= \begin{cases} \frac{(b-t)^3(2b+a-3t)}{12} & \text{si } t \leq \frac{a+b}{2}, \\ \frac{(a-t)^3(b+2a-3t)}{12} & \text{si } t \geq \frac{a+b}{2}. \end{cases}$$

Ainsi, comme le noyau K est négatif,

$$E(f) = -f^{(4)}(\xi)\frac{(b-a)^5}{2880} \text{ pour un certain } \xi \in [a,b].$$

3. La formule d'intégration approchée de Gauss

On rappelle du Lemme 3.3 que toute méthode d'intégration approchée est au plus d'ordre $2n+1$. On peut alors se demander s'il existe une méthode d'intégration approchée qui est exactement d'ordre $m = 2n+1$. La réponse à cette question est oui, si on choisit bien les noeuds $x_0, \ldots, x_n$.

Théorème 3.12 (Gauss). *Soit $[a,b] = [-1,1]$, et soit $n \geq 0$. Alors il existe des noeuds $a \leq x_0 < \cdots < x_n \leq b$ et une méthode d'intégration approchée $J(f) = \sum_{i=0}^n \lambda_i f(x_i)$ qui est d'ordre $2n+1$. Plus précisément, il suffit de prendre comme noeuds les racines du $n+1$-ième polynôme de Legendre L_{n+1} et comme méthode d'intégration approchée celle du Théorème 3.1.*

Démonstration. Soit L_{n+1} le $n+1$-ième polynôme de Legendre. On rappelle que la suite des polynômes de Legendre est obtenue par le procédé de Gram-Schmidt appliqué à la suite des monômes, en utilisant le produit scalaire $\langle f, g\rangle = \int_{-1}^1 f(x)g(x)\,dx$. Le polynôme L_{n+1} est un polynôme de degré $n+1$ qui est orthogonal à tous les polynômes de degré $\leq n$ (Lemme 2.11). En plus, toutes les racines de L_{n+1} sont simples et contenues dans l'intervalle $[-1,1]$ (Théorème 2.12)

Soit maintenant p un polynôme de degré $\leq 2n+1$. Une division avec reste montre que

$$p = q\,L_{n+1} + r$$

pour des polynômes de degré $\leq n$. Alors

$$\begin{aligned}
I(p) &= \int_{-1}^{1} p(x)\,dx \\
&= \int_{-1}^{1} q(x)L_{n+1}(x)\,dx + \int_{-1}^{1} r(x)\,dx \\
&= \int_{-1}^{1} r(x)\,dx && (q \text{ et } L_{n+1} \text{ sont orthogonaux}) \\
&= \sum_{i=0}^{n} \lambda_i\, r(x_i) && (J \text{ est d'ordre } m \geq n) \\
&= \sum_{i=0}^{n} \lambda_i\, q(x_i)L_{n+1}(x_i) + \sum_{i=0}^{n} \lambda_i\, r(x_i) && (x_i \text{ sont racines de } L_{n+1}) \\
&= \sum_{i=0}^{n} \lambda_i\, p(x_i) \\
&= J(p).
\end{aligned}$$

Comme $p \in \mathbb{R}_{2n+1}[X]$ était arbitraire, ceci montre que J est d'ordre $m \geq 2n+1$. Par le Lemme 3.3, $m \leq 2n+1$, et donc l'ordre $m = 2n+1$. □

4. Différentiation numérique

Etant donné une fonction $f \in C^1([a,b])$ et un point $x \in [a,b]$, on souhaite calculer la dérivée $L(f) = f'(x)$. On souhaite calculer cette dérivée à l'aide d'une formule à deux noeuds

$$\Lambda(f) = \lambda_0 f(a) + \lambda_1 f(b),$$

où λ_0, λ_1 sont des constantes. On veut que la formule $\Lambda(f)$ soit exacte pour tous les polynômes de degré ≤ 1. On choisissant $f(x) = 1$ et $f(x) = x$ on trouve alors les deux conditions

$$\begin{aligned}
\lambda_0 + \lambda_1 &= 0 \quad \text{et} \\
\lambda_0\, a + \lambda_1\, b &= 1,
\end{aligned}$$

et donc $-\lambda_0 = \lambda_1 = \frac{1}{b-a}$. Ainsi

$$\Lambda(f) = \frac{f(b) - f(a)}{b-a}.$$

4.1. Schéma centré. Ici, on prend $x = \frac{a+b}{2}$, c.à.d. on souhaite calculer la dérivée $L(f) = f'(\frac{a+b}{2})$ au milieu de l'intervalle $[a,b]$. Par construction, l'erreur

$$E(f) = L(f) - \Lambda(f) = f'(\frac{a+b}{2}) - \frac{f(b) - f(a)}{b-a}$$

est nulle pour tout polynôme f de degré ≤ 1. Mais on voit facilement que $E(x^2) = 0$ aussi, et que $E(x^3) \neq 0$. On dit que la méthode de différentiation $\Lambda(f)$ est d'ordre 2.

Afin d'estimer l'erreur commise, lorsque $f \in C^3([a,b])$, on utilise la formule de Taylor (Théorème 3.7):

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \int_a^b \frac{(x-t)_+^2}{2} f^{(3)}(t)\, dt,$$

où

$$t_+ = \begin{cases} t & \text{si } t \geq 0, \\ 0 & \text{si } t < 0. \end{cases}$$

Puisque la formule de différentiation $\Lambda(f)$ est d'ordre 2, on trouve

$$\begin{aligned} E(f) &= E(x \mapsto \int_a^b \frac{(x-t)_+^2}{2} f^{(3)}(t)\, dt) \\ &= \frac{d}{dx}\int_a^b \frac{(x-t)_+^2}{2} f^{(3)}(t)\, dt\Big|_{x=\frac{a+b}{2}} - \\ &\quad - \frac{1}{b-a}\Big(\int_a^b \frac{(b-t)_+^2}{2} f^{(3)}(t)\, dt - \int_a^b \frac{(a-t)_+^2}{2} f^{(3)}(t)\, dt\Big) \\ &= \int_a^b (x-t)_+ f^{(3)}(t)\, dt\Big|_{x=\frac{a+b}{2}} - \frac{1}{b-a}\int_a^b \frac{(b-t)^2}{2} f^{(3)}(t)\, dt \\ &= \int_a^b K(t) f^{(3)}(t)\, dt \end{aligned}$$

avec

$$K(t) = \begin{cases} -\frac{(a-t)^2}{2(b-a)} & \text{si } t \leq \frac{a+b}{2}, \\ -\frac{(b-t)^2}{2(b-a)} & \text{si } t \geq \frac{a+b}{2}. \end{cases}.$$

Ainsi,

$$E(f) = f^{(3)}(\xi)\int_a^b K(t)\, dt = -f^{(3)}(\xi)\frac{(b-a)^2}{24},$$

ou

$$|E(f)| \leq \|f^{(3)}\|_\infty \frac{(b-a)^2}{24}.$$

4.2. Schéma décentré. Ici, $x \in [a,b]$ est arbitraire. Un simple calcul montre que $E(f) = f'(x) - \frac{f(b)-f(a)}{b-a} = 0$ pour $f(x) = x^2$ si et seulement si $x = \frac{a+b}{2}$ (schéma centré), c.à.d. la formule de différentiation $\Lambda(f)$ est d'ordre 2 si et seulement si $x = \frac{a+b}{2}$. Dans le cas $x \neq \frac{a+b}{2}$, elle est d'ordre 1.

Afin d'estimer l'erreur commise dans ce cas et pour une fonction $f \in C^2([a,b])$, on utilise la formule de Taylor (Théorème 3.7)

$$f(x) = f(a) + f'(a)(x-a) + \int_a^b (x-t)_+ f''(t)\, dt.$$

Puisque la formule de différentiation $\Lambda(f)$ est d'ordre 1, on trouve

$$\begin{aligned} E(f) &= E(x \mapsto \int_a^b (x-t)_+ f''(t)\, dt) \\ &= \frac{d}{dx} \int_a^b (x-t)_+ f''(t)\, dt \Big|_x - \\ &\quad - \frac{1}{b-a}\Big(\int_a^b (b-t)_+ f''(t)\, dt - \int_a^b (a-t)_+ f''(t)\, dt\Big) \\ &= \int_x^b f''(t)\, dt - \frac{1}{b-a} \int_a^b (b-t) f''(t)\, dt \\ &= \int_a^b K(t) f''(t)\, dt \end{aligned}$$

avec

$$K(t) = \begin{cases} \frac{t-a}{b-a} & \text{si } t \le x, \\ \frac{t-b}{b-a} & \text{si } t \ge \frac{a+b}{2}. \end{cases}.$$

Dans ce cas on a l'estimation

$$|E(f)| \le \|f''\|_\infty \int_a^b |K(t)|\, dt = \|f''\|_\infty \frac{b-a}{2}.$$

CHAPITRE 4

Résolution numérique d'équations non-linéaires dans $\mathbb{R}^N$

Dans ce chapitre on s'intéresse à résoudre l'équation

$$f(x) = y$$

pour une fonction $f : \mathbb{R}^N \to \mathbb{R}^N$ donnée et un vecteur $y \in \mathbb{R}^N$ donné, ou l'équation

$$g(x) = x$$

pour une fonction $g : \mathbb{R}^N \to \mathbb{R}^N$ donnée (en fait, les fonctions f et g peuvent être définies sur un domaine de définition inclus dans $\mathbb{R}^N$).

1. Résolution numérique d'une équation d'une variable par dichotomie ou regula falsi

On rappelle le résultat suivant d'analyse des fonctions d'une variable réelle.

THÉORÈME 4.1 (Théorème de la valeur intermédiaire). *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue sur un intervalle compact. Alors pour tout $f(a) \leq \eta \leq f(b)$ (ou $f(b) \leq \eta \leq f(a)$, si $f(b) \leq f(a)$) il existe un $c \in [a, b]$ tel que $f(c) = \eta$.*

Comme corollaire on obtient immédiatement qu'une fonction continue $f : [a, b] \to \mathbb{R}$ vérifiant $f(a)f(b) \leq 0$ admet une solution $x \in [a, b]$ de l'équation $f(x) = 0$. Ceci est une motivation pour l'algorithme suivant.

Algorithme. On suppose que $f : [a, b] \to \mathbb{R}$ est une fonction continue telle que $f(a)f(b) \leq 0$.
$1^{ère}$ *étape.* On pose $a_0 = a$ et $b_0 = b$. Alors $f(a_0)f(b_0) \leq 0$ (par hypothèse sur f). Si $f(a_0) = 0$ ou $f(b_0) = 0$, alors on a déjà trouvé une solution de l'équation $f(x) = 0$ et l'algorithme s'arrête. Sinon, on a $f(a_0)\, f(b_0) < 0$ et on continue avec la deuxième étape.
$2^{ème}$ *étape.* Soit $k \geq 0$. On suppose que $f(a_k)\, f(b_k) < 0$ et on choisit $c_k \in]a_k, b_k[$. Alors exactement un des trois cas suivants est vérifié.
Cas 1: $f(c_k) = 0$.
Cas 2: $f(a_k)\, f(c_k) < 0$.
Cas 3: $f(c_k)\, f(b_k) < 0$.
Si le premier cas est vérifié, alors $x = c_k$ est une solution du problème $f(x) = 0$ et l'algorithme s'arrête. Sinon, dans le deuxième cas on pose $a_{k+1} = a_k$ et $b_{k+1} = c_k$. Dans le troisième cas, on pose $a_{k+1} = c_k$ et $b_{k+1} = b_k$. Avec ces définitions on continue avec la deuxième étape, en remplaçant k par $k + 1$.

Soit cet algorithme s'arrête en un nombre fini d'étapes (et dans ce cas on a trouvé une solution du problème $f(x) = 0$), soit cet algorithme construit trois suites (a_k), (b_k) et (x_k) telles que

(i) $\lim_{k\to\infty} a_k$ existe (car (a_k) est une suite croissante),
(ii) $\lim_{k\to\infty} b_k$ existe (car (b_k) est une suite décroissante),
(iii) quelque soit $k \geq 0$, l'intervalle $[a_k, b_k]$ contient une solution $x = x_k$ du problème $f(x) = 0$.

Si

$$\lim_{k\to\infty} b_k - a_k = 0,$$

alors

$$x := \lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k = \lim_{k\to\infty} x_k \text{ et } f(x) = 0.$$

Dans ce cas, l'algorithme converge donc vers une solution du problème $f(x) = 0$.

Exemple 4.2. (a) A chaque étape on choisit

$$c_k = \frac{a_k + b_k}{2}.$$

Dans cet exemple on a

$$b_k - a_k = 2^{-k}(b - a)$$

et donc $\lim_{k\to\infty} b_k - a_k = 0$. Dans cet exemple, l'algorithme converge vers une solution du problème $f(x) = 0$.

(b)

2. Approximations successives

Théorème 4.3 (Théorème du point fixe de Banach).

3. Méthode de Newton

4. Méthode de plus grande descente

CHAPITRE 5

Résolution numérique d'équations différentielles ordinaires

Dans ce chapitre, soit $f : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$, $f = f(t, x)$, une fonction continue donnée, soit $t_0 \in \mathbb{R}$ un temps initial et $x_0 \in \mathbb{R}^N$ une donnée initiale. On cherche à résoudre l'équation différentielle ordinaire

$$\begin{cases} \dot{x}(t) = f(t, x(t)) \\ x(t_0) = x_0. \end{cases} \tag{5.1}$$

On rappelle le thèoreme d'existence locale de Peano.

THÉORÈME 5.1 (Peano). *L'équation différentielle* (5.1) *admet toujours une* solution locale*, c.à.d. il existe un intervalle $I \subseteq \mathbb{R}$ qui est un voisinage de t_0, et il existe une fonction $x : I \to \mathbb{R}^N$ de classe C^1 telle que $\dot{x}(t) = f(t, x(t))$ pour tout $t \in I$ et $x(t_0) = x_0$.*

DÉMONSTRATION CONSTRUCTIVE. On choisit $\tau > 0$ et $r > 0$. Comme f est continue, on a

$$M := \sup_{\substack{|t-t_0| \le \tau \\ \|x-x_0\| \le r}} \|f(t, x)\| < +\infty.$$

En choisissant $\tau > 0$ plus petit, si nécessaire, on peut supposer que $\tau M \le r$.

Soit $n \ge 1$. On pose $t_k := t_0 + \frac{k\tau}{n}$ $(k = 0, \dots, n)$. Ensuite, on définit une fonction $x_n : [t_0, t_0 + \tau] \to \mathbb{R}^N$ en posant

$$\begin{aligned} x_n(t_0) &= x_0 \qquad \text{et} \\ x_n(t_{k+1}) &= x_n(t_k) + \frac{\tau}{n} f(t_k, x_n(t_k)) \quad \text{pour } k = 0, \dots, n-1, \end{aligned} \tag{5.2}$$

et en prolongeant x_n en une fonction affine par morceaux, c.à.d. affine sur les intervalles $[t_k, t_{k+1}]$:

$$\begin{aligned} x_n(t) &= x_n(t_k) + (x_n(t_{k+1}) - x_n(t_k))(t - t_0 - \frac{k\tau}{n}) \\ &= x_n(t_k) + \frac{\tau}{n} f(t_k, x_n(t_k))(t - t_0) \quad \text{si } t \in [t_k, t_{k+1}]. \end{aligned} \tag{5.3}$$

On montre que pour tout $n \ge 1$ et tout $k = 0, \dots, n$, on a

$$\|x_n(t_k) - x_0\| \le \frac{k}{n} r. \tag{5.4}$$

Pour cela, on fixe $n \geq 1$. Par définition de x_n, l'estimation (5.4) est vraie pour $k = 0$. On suppose alors qu'elle est vraie pour un $k \in \{0, \dots, n-1\}$. Alors

$$\begin{aligned} \|x_n(t_{k+1}) - x_0\| &\leq \|x_n(t_k) - x_0\| + \|\frac{\tau}{n} f(t_k, x_n(t_k))\| \\ &\leq \frac{k}{n}\, r + \frac{\tau}{n}\, M \\ &\leq \frac{k}{n}\, r + \frac{1}{n}\, r \\ &\leq \frac{k+1}{n}\, r. \end{aligned}$$

On a donc démontré (5.4) par récurrence.

De l'estimation (5.4) on obtient que pour tout $n \geq 1$ et tout $t \in [t_0, t_0 + \tau]$

$$|t - t_0| \leq \tau,\ \|x(t) - x_0\| \leq r \text{ et } \|f(t, x(t))\| \leq M.$$

En particulier, la suite (x_n) est bornée dans $C([t_0, t_0 + \tau]; \mathbb{R}^N)$. On montre ensuite que pour tout $n \geq 1$ et tout $s,\ t \in [t_0, t_0 + \tau]$ $(s \leq t)$ on a

$$x_n(t) = x_n(s) + \int_s^t \hat{x}_n(r)\ dr,$$

où

$$\hat{x}_n(t) = f(t_k, x_n(t_k)) \quad \text{si } t \in [t_k, t_{k+1}].$$

Ainsi,

$$\|x_n(t) - x_n(s)\| \leq M\,|t - s| \quad \text{pour tout } s,\ t \in [t_0, t_0 + \tau],$$

c.à.d. les x_n sont lipschitziennes avec une constante de Lipschitz qui ne dépend pas de n.

On montre finalement que (x_n) admet une sous-suite convergente et que la limite x de cette sous-suite est une solution de l'équation différentielle (5.1). □

1. Le schéma d'Euler explicite ou implicite

Le schéma (5.2) est appelé le *schéma d'Euler explicite*. L'idée de ce schéma est simple: on remplace dans l'équation différentielle $\dot{x}(t) = f(t, x(t))$ la dérivée $\dot{x}$ en un point t_k par la différence finie $\frac{n}{\tau}\,(x_n(t_{k+1}) - x_n(t_k))$, et on résoud l'équation

$$\frac{n}{\tau}\,(x_n(t_{k+1}) - x_n(t_k)) = f(t_k, x_n(t_k))$$

de manière récursive. La démonstration du Théorème de Peano montre que les solutions approchées x_n ainsi construites convergent uniformément (après passage à une sous-suite) vers une solution exacte x de l'équation différentielle (5.1).

On note, sans démonstration, que l'on peut remplacer dans la définition des solutions approchées les subdivisions équi-distantes par des subdivisions quelconques

$\sigma : t_0 < t_1 < \cdots < t_n = t_0 + \tau$ de l'intervalle $[t_0, t_0 + \tau]$. Pour une telle subdivision σ la solution approchée associée est définie par

$$\begin{aligned} x_\sigma(t_0) &= x_0 \qquad \text{et} \\ x_\sigma(t_{k+1}) &= x_\sigma(t_k) + h_k f(t_k, x_\sigma(t_k)) \quad \text{pour } k = 0, \dots, n-1, \end{aligned} \tag{5.5}$$

où h_k est le pas :

$$h_k := t_{k+1} - t_k \quad (k = 0, \dots, n-1).$$

Le schéma (5.5) est aussi appelé *schéma d'Euler explicite*. On appelle $|\sigma| := \sup_k h_k$ la *norme* de la subdivision σ. Alors (x_σ) admet une "sous-suite" qui converge (lorsque $|\sigma| \to 0$) vers une solution x du problème (5.1).

On peut motiver le schéma d'Euler aussi en remarquant que pour toute solution x du problème (5.1) et toute subdivision $\sigma : t_0 < t_1 < \cdots < t_n = \tau$ de l'intervalle $[t_0, t_0 + \tau]$ on a

$$\begin{aligned} x(t_0) &= x_0 \qquad \text{et} \\ x(t_{k+1}) &= x(t_k) + \int_{t_k}^{t_{k+1}} f(t, x(t))\, dt \quad \text{pour } k = 0, \dots, n-1. \end{aligned} \tag{5.6}$$

Ensuite, en remplaçant l'intégrale $\int_{t_k}^{t_{k+1}} f(t, x(t))\ dt$ par $h_k\, f(t_k, x(t_k))$ (méthode des rectangles à gauche), on trouve le schéma d'Euler explicite (5.5). Si on remplace l'intégrale $\int_{t_k}^{t_{k+1}} f(t, x(t))\ dt$ par $h_k\, f(t_{k+1}, x(t_{k+1}))$ (méthode des rectangles à droite), on trouve le *schéma d'Euler implicite* :

$$\begin{aligned} x_\sigma(t_0) &= x_0 \qquad \text{et} \\ x_\sigma(t_{k+1}) &= x_\sigma(t_k) + h_k f(t_{k+1}, x_\sigma(t_{k+1})) \quad \text{pour } k = 0, \dots, n-1. \end{aligned} \tag{5.7}$$

Afin de trouver $x_\sigma(t_{k+1})$ dans ce schéma implicite, il faut dans chaque itération résoudre une équation non-linéaire. Ceci est à première vue plus difficile que seulement évaluer $f(t_k, x_\sigma(t_k))$, mais il y a des situations où il est préferable d'utiliser un schéma implicite.

Estimation de l'erreur. Dans la suite, on fixe une subdivision $\sigma : t_0 < t_1 < \cdots < t_n = t_0 + \tau$ et une solution approchée x_σ associée, donnée par le schéma d'Euler explicite (5.5). Soit x une solution exacte du problème (5.1). On définit *l'erreur de discrétisation*

$$e_k := x(t_k) - x_\sigma(t_k). \tag{5.8}$$

On veut estimer cet erreur en fonction de la norme

$$h := |\sigma| = \sup_{0 \le k \le n-1} h_k = 0.$$

Nous supposons que la fonction f est localement lipschitzienne par rapport à la deuxième variable, c.à.d. qu'il existe une constante $L \ge 0$ telle que pour tout $t \in [t_0, t_0 + \tau]$ et tout $x, y \in \mathbb{R}^N$ avec $\|x - x_0\| \le r$ et $\|y - x_0\| \le r$ on a

$$\|f(t, x) - f(t, y)\| \le L\, \|x - y\|. \tag{5.9}$$

REMARQUE 5.2. Sous l'hypothèse (5.9) on peut montrer que la solution locale x construit dans le Théorème de Peano est (localement) unique dans le sens que si $y : [t_0, t_0 + \tau'] \to \mathbb{R}^N$ est une deuxième solution de (5.1) et si $\tau' \leq \tau$, alors $y(t) = x(t)$ pour tout $t \in [t_0, t_0 + \tau']$ (Théorème de Cauchy-Lipschitz).

On définit aussi *l'erreur de troncature*

$$\begin{aligned} \varepsilon_{k+1} &:= \frac{1}{h_k}\left(x(t_{k+1}) - x(t_k) - h_k f(t_k, x(t_k))\right) \\ &= \frac{1}{h_k}\Big(\int_{t_k}^{t_{k+1}} f(t, x(t))\, dt - h_k f(t_k, x(t_k))\Big). \end{aligned} \tag{5.10}$$

Finalement, étan donné une fonction $g \in C([t_0, t_0+\tau]; \mathbb{R}^N)$ et une constante $\delta > 0$, on définit le *module de continuité*

$$\omega(g, \delta) := \max_{\substack{t,s\in[t_0,t_0+\tau] \\ |t-s|\leq\delta}} \|g(t) - g(s)\|.$$

Avec cette définition, l'erreur de troncature peut être majoré par

$$|\varepsilon_{k+1}| = \Big|\frac{1}{h_k}\int_{t_k}^{t_{k+1}} (\dot{x}(t) - \dot{x}(t_k))\, dt\Big| \leq \omega(\dot{x}, |\sigma|), \quad k = 0, \ldots, n-1.$$

Comme $\dot{x}$ est uniformément continue, on a $\lim_{|\sigma|\to 0} \omega(\dot{x}, |\sigma|) = 0$. Supposons en plus que f est de classe C^1. Alors la solution x de (5.1) est de classe C^2 et on a

$$\ddot{x}(t) = \frac{\partial f}{\partial t}(t, x(t)) + \frac{\partial f}{\partial x}(t, x(t)) f(t, x(t)).$$

En utilisant la formule de Taylor, on en déduit

$$|\varepsilon_{k+1}| = \Big|\frac{1}{h_k}\int_{t_k} t_{k+1}(t_{k+1} - t)\ddot{x}(t)\, dt\Big| \leq \int_{t_k}^{t_{k+1}} |\ddot{x}(t)|\, dt.$$

Ainsi,

$$e_{k+1} = e_k + h_k(f(t_k, y(t_k)) - f(t_k, x_\sigma(t_k))) + h_k\varepsilon_{k+1}.$$

Ceci implique, comme f est lipschitzienne,

$$\|e_k\| \leq$$